January 1997

Behavioral and Judged Coordination in Adult Informal Social Interactions: Vocal and Kinesic Indicators

Joseph N. Cappella
University of Pennsylvania, jcappella@asc.upenn.edu

Recommended Citation

Cappella, J. N. (1997). Behavioral and Judged Coordination in Adult Informal Social Interactions: Vocal and Kinesic Indicators. *Journal of Personality and Social Psychology, 72* (1), 119-131. https://doi.org/10.1037/0022-3514.72.1.119

Behavioral and Judged Coordination in Adult Informal Social Interactions: Vocal and Kinesic Indicators

Abstract
Coordination in social interaction means that persons adjust their actions to those of their partners. Common methods for measuring coordination include judgments and behavioral covariation. Sixteen 1-min segments of interaction were chosen (8 high and 8 low in behavioral coordination). In Study 1, 51 people judged the 16 segments, rating each for coordination. Study 2 (N = 17) used different items. Study 3 (N = 22) replicated Study 2 without sound and with a mosaic pattern imposed on the faces. Results indicated judges were reliable, able to distinguish high from low coordination interactions on the basis of 1-min slices for male but not female dyads. Segments judged to be coordinated had partners smiling in synchrony but with complementary patterns of gazing and gesturing. Both measures correlated with conversational satisfaction, but only behavioral coordination predicted attraction.

Journal of Personality and Social Psychology
1997, Vol. 72, No. 1, 119–131

0022-3514/97/$3.00

Behavioral and Judged Coordination in Adult Informal Social Interactions: Vocal and Kinesic Indicators

Joseph N. Cappella
University of Pennsylvania

Coordination in social interaction means that persons adjust their actions to those of their partners. Common methods for measuring coordination include judgments and behavioral covariation. Sixteen 1-min segments of interaction were chosen (8 high and 8 low in behavioral coordination). In Study 1, 51 people judged the 16 segments, rating each for coordination. Study 2 ($N = 17$) used different items. Study 3 ($N = 22$) replicated Study 2 without sound and with a mosaic pattern imposed on the faces. Results indicated judges were reliable, able to distinguish high from low coordination interactions on the basis of 1-min slices for male but not female dyads. Segments judged to be coordinated had partners smiling in synchrony but with complementary patterns of gazing and gesturing. Both measures correlated with conversational satisfaction, but only behavioral coordination predicted attraction.

The success or failure of personal relationships is often dependent on the nature of the interaction that the partners enact. Their personal styles may clash or be harmonious; they may feel like they are "in synch" with one another or at odds. Metaphors like synchrony, harmony, meshing, coordination, and so on imply that good relationships result from interactions in which there is fit between partners. These and other metaphors describing concordant and discordant interactional styles have had a long history, dating at least to the empirical and mathematical work of anthropologist Eliot Chapple (1940). Until relatively recently, the metaphors have remained richer than the scientific models spawned by them.

One important theory driving recent research concerned with the relationship between interaction and interpersonal outcomes is the coordination–rapport hypothesis (Tickle-Degnen & Rosenthal, 1987). This hypothesis suggests that various interpersonal outcomes including (but not limited to) attraction, satisfaction, attachment, longevity, and rapport covary with the kind and intensity of interactional patterns enacted by partners. In particular, coordinated patterns should be preferred by partners more than uncoordinated patterns. Coordinated patterns of interaction are ones that are similar in tempo, matched in position and movement, in synch with one another, and meshed. Bernieri and Rosenthal (1991) described behavioral coordination as including the matching of behaviors, the adoption of similar behavioral rhythms, the manifestation of simultaneous movement, and the "interrelatedness" of individual behaviors. No one of these measures is definitive, nor the set necessarily exhaustive.

Coordination does not imply that interactions are necessarily reciprocal or complementary, but it does imply that partners are responsive to one another's actions. The essence of the definition of coordination, indeed, the essence of the definition of interpersonal communication, is the notion that partners are mutually responsive. Both flirting and fighting require partners to be responsive to one another's actions.

Although the conceptual center of coordination as a construct is clear enough, its operationalization has been too diverse to permit scientific knowledge on coordination and rapport to accumulate easily. Bernieri and Rosenthal (1991) described two broad operational approaches: *behavioral ratings* (also called *behavioral judgments* elsewhere by Rosenthal, 1987) and *behavioral coding* (also called *microanalysis*). The purpose of the trio of studies reported here is to compare judges' evaluations of coordination with those provided through behavioral coding.

Research on Coordination and Interpersonal Outcome

Extended reviews of the literature relating coordination to rapport are available elsewhere (Bernieri & Rosenthal, 1991; Cappella, 1994), but highlighting a few studies will give a sense of the nature of the findings that reside under the umbrella of coordination-rapport. Among married couples, the degree of reciprocity in hostile affect covaries with marital satisfaction. Gottman's (1979) widely cited findings show that although all of his couples tend to show reciprocity in hostile affect in discussions about common problems in their marriages, the less well-adjusted showed greater reciprocity in hostile affect than did the better adjusted couples. These findings have been replicated by Margolin and Wampold (1981). Pike and Sillars (1985) also found greater reciprocity in negative vocal affect for dissatisfied as opposed to satisfied married couples.

Interactional coordination is consequential for attachment between infants and their caretakers. Isabella, Belsky, and van Eye (1989) tested mothers and their infants. The infants were

I thank Frank Bernieri of the Department of Psychology at University of Toledo, who gathered the data for Studies 2 and 3 reported here and kindly made them available for analysis. The conclusions and interpretations from these data are mine alone. An earlier version of this article was presented at the International Communication Association meetings, Chicago, May 1996.

Correspondence concerning this article should be addressed to Joseph N. Cappella, Annenberg School for Communication, 3620 Walnut Street, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6220. Electronic mail may be sent via the Internet to fjc@asc.upenn.edu.

observed at 1, 3, and 9 months of age interacting with their mothers and were categorized as primarily synchronous (coordinated) or asynchronous (noncoordinated). These pairs were later tested and found to be secure, avoidant, or resistant at age 1 year according to the criteria of Ainsworth's Strange Situation. Pairs that were coordinated at ages 1 and 3 months were also securely attached at 1 year. Isabella and Belsky (1991) have replicated these findings, adding to the robustness of the link between the interactional sensitivity of the mother and her infant and the quality of their attachment.

Bernieri, Reznick, and Rosenthal (1988) also studied infants interacting with their mothers and with strangers. They created real and pseudo-interactions. The pseudo-interactions involved juxtaposing video images of an infant with the mother that were taken from different segments of their actual interaction. Judges evaluated brief segments of the various types of interactions on criteria related to interactional coordination. The judges rated real mother–infant segments as more coordinated than the pseudo-interactions. Judges' ratings of coordination also correlated positively with a separate sample's judgments of positive affect expressed in the segments.

Using the same judgment technique with an adult population, Bernieri (1988) found that pairs of unacquainted students who were judged to be more behaviorally coordinated also reported higher levels of positive rapport with one another. This finding remained even after the degree of friendliness, activity, and attention (mean evaluations provided by a separate sample) were partialled out of the correlation between rapport and judged coordination.

Attraction between unacquainted partners has also been shown to covary with coordination on various nonverbal behaviors (Cappella & Flagg, 1992; Cappella, Palmer, & Donzella, 1991) as well. In these studies, coordination was defined through behavioral coding rather than by judgment techniques. Time series measures of association between Person A's and Partner B's behavioral series were the operational measures of coordination. The greater the degree of association between the A series and the B series, the more attracted the partners were to one another (this finding did not hold for all types of partners).

The research literature shows that coordination in interaction is associated with important interpersonal outcomes in both adult and infant–adult social interactions.

Measures of Coordination

The studies of coordination and rapport just discussed seem to be testing the same hypothesis. However, they use very different definitions of coordination. The studies by Bernieri and his colleagues operationally define coordination in terms of judges' ratings of brief segments of interaction for the presence of interactional synchrony. I will refer to this as the *judgment method.* The other studies code specific verbal and nonverbal behaviors, array the behaviors in sequence for each partner, and calculate the degree and magnitude of covariation between partners' behaviors using some form of time series analysis. I will call this the *behavioral coding method.*

The judgment method is less expensive than the coding method, which requires evaluations of brief segments of interaction by groups of judges who are not trained. Entire interactions do not need to be used to obtain reasonably representative judgments of the partners' interactional behavior (Ambady & Rosenthal, 1992). The judgments also appear to be valid indicators of something that is interpersonally important. Otherwise the judgments would not be so clearly related to assessments of affect nor to partners' own reports of attraction.

What is not clear is what the judgments are based on. Are judges focusing on specific behaviors and excluding others? Are they making gestalt evaluations attributable to general patterns but not to specific behaviors? Are they really using criteria related to temporal matching, mismatching, and covariation in partners' behaviors, or are they using simpler heuristics, ignoring covariation, and, for example, estimating simple, static behavioral activity levels? In this last case, behavioral measures of covariation between partners would not correlate with judges' scores, but durational or frequency measures of individuals' behaviors would.

The coding method is costly in the time required for training, in the number of behaviors coded, and in the assumed necessity of coding long periods of interaction. Like the judgment method, the coding method appears to capture something that is interpersonally consequential and that covaries with marital adjustment, infant attachment, and attraction in unacquainted pairs. On the other hand, the microscopic aspects of coding may be too fine-grained to matter to human observers. Indexes of behavioral coordination may not be sophisticated enough to capture the (potentially) complex ways that human observers weigh the mix of behaviors in creating a judgment of synchronous and asynchronous interaction.

The question asked in this article concerns the relationship between the coding and judgment methods of operationalizing coordination. If the two methods are unrelated, then it is not clear what they are measuring. The judgment method asks raters to assess aspects of interaction such as similarity of tempo or movement coordination. If the judgments are unrelated to behavioral measures of tempo or movement synchrony, then it would be difficult to claim that the construct being operationalized is actually behaviorally based.

If the behavioral coding method and judged coordination fail to correlate, then one might conclude that observers are not sensitive to behavioral coordination. Of course, they may be seeing some other facet of coordination not tapped by behavioral coding, but what that facet is would continue to be hidden.

If the judgment and coding methods are positively correlated, then several other questions arise immediately. The first is what behaviors account for judgments of interactional coordination? The second concerns the predictive ability of the measures in terms of shared variance and in predicting outcomes (such as reported attraction). Answers to the first question will yield clues to what behaviors observers use in making judgments of interactional coordination and, by implication, what behaviors are important to judgments of interpersonal outcome. Answers to the second question will indicate whether the coding and judging methods are interchangeable.

Nonverbal Behaviors and Coordination

Behaviors

In this study, only vocal and kinesic behaviors are examined. Other specific behaviors, such as verbal ones, may be highly relevant; too, very general patterns of coordination may lie at the heart of the judged coordination construct. This study assesses neither verbal nor gestalt aspects of coordination, even though the former can be central to the judgments persons make about one another during interaction (Palmer, 1989; van Lear, 1991), and the latter to a part of the perceptual filter of observers (Baron, Amazeen, & Beek, 1994; Newtson, 1994).

Two categories of nonverbal behavior are examined: behaviors that indicate activation (sometimes called *involvement* or *animation*) and behaviors associated with turn-taking (Duncan & Fiske, 1977; Patterson, 1983). These two categories were chosen because they are most likely to be relevant to judged coordination.

Two components of judged coordination used in previous research are tempo matching and behavioral meshing. The former suggests that periods of high and low activity are matched by partners' periods of high and low activity. The latter implies that partners mesh when they avoid transgressing on one another's interactional space. That is, they should avoid overlapping one another's turns. Judged coordination may be associated with similarity in behavioral tempo between partners and complementarity in turn-taking behaviors.

Judged coordination may be lowered when partners interrupt one another or overlap one another's conversational turns. Judged coordination may be higher when turns remain distinct. This effect would manifest itself as a negative correlation between judged coordination and a measure of covariation between partners' floor-holding behaviors.

Judged coordination may be elevated when partners' behavioral activity levels rise together and fall together; the coordination may be lowered when activation levels are complementary. This effect would manifest itself as a positive correlation between judged coordination and a measure of covariation between partners' involvement behaviors.

Despite the language suggesting specific hypotheses linking judged and behavioral coordination, how particular behaviors will function is less clear. The problem is that the same behavior may both be an indicator of turn-taking and an indicator of activation. For example, consider illustrative gestures—those that accompany speech (McNeill, 1992). These gestures are associated with holding the floor, and their absence with being in the listening role. At the same time, this behavior can be seen as a sign of animation or involvement.

When one partner gestures a lot and the other only a little and vice versa, their coordination on gesture is negative. But illustrative gestures only occur during speech. Thus, in the absence of speech overlap, neither person may gesture at the same time as the partner. If they never gesture simultaneously, judges might evaluate them as coordinated in the sense that their gestures mesh—neither intrudes on the conversational space of the other. With one partner high and one very low in gestural animation over a segment of time, judges might rate them as uncoordinated in tempo—one is high and one is low in a behavior indicating animation.

The point is that it is not always clear how judges will use a particular behavior in making judgments about coordination when the behavior is an indicator of both turn-taking and animation. The problem with illustrative gestures is also true of face-directed gaze, vocalization, and to a lesser extent, adaptor gestures (hand movements directed at one's body, clothing, and artifacts). Each of these behaviors is associated with holding the floor. Gaze is averted more while speaking than while listening. Vocalization is greater while holding the floor (in fact, it defines holding the floor) than while in the listening role. Adaptor gestures are more common while speaking than while listening. Although associated with speaking–listening roles, each of the behaviors is also a sign of behavioral involvement. The greater the gaze, gesture, and vocalization, the more activated the person is seen to be (Cappella & Street, 1985).

Unlike the preceding four behaviors, which could function as signs of animation or as signs of speaker–hearer role, smiles do not have an appreciable correlation with conversational role. Smiles occur with roughly equal likelihood by listeners and speakers. They are also typically a sign of involvement (and positive affect). Thus, we expect smiles to function solely as an involvement behavior, with judged coordination correlating positively with partner covariation in smiling.

Aside from smiling, no directional hypotheses can be offered about correlations between vocal coordination and judged coordination; gaze coordination and judged coordination; or gestural coordination (either illustrative or adaptor) and judged coordination.

Operational Measures

Operational measures of behavioral coordination are numerous. The assumption adopted here is that the only acceptable measures of behavioral coordination are ones that take into account temporal covariation between partners. Omitted are aggregate measures of similarity such as mean difference. Partners are coordinated behaviorally when there is some covariation over time between their behaviors.

In the studies reported here, simple contemporaneous correlation is the usual measure of behavioral coordination. Fancier time series measures would seem to require much more complex evaluations by judges than would simple correlation. The bet is that people are judging coordination through simple correlation between partners' behavior over time.

Predicting Outcomes

The final hypotheses to be tested assess the predictive abilities of the two measures of coordination. The variance explained in judged coordination on the basis of behavioral coordination will indicate the degree of interchangeability of the two measures. High levels of common variance will allow the cheaper measures like judged coordination to replace their more expensive behavioral counterparts. Even if shared variance is too low to allow interchangeability, knowing which elements of behavioral coordination predict judged coordination will extend the claims that can be made about the implications of high and low levels of judged coordination.

A second kind of predictability concerns which method is more successful in predicting interpersonal outcomes, in this case the partners' attraction to one another and their attitudes toward the conversation itself. This criterion is especially important because it is directly tied to the consequences of the interaction for the participants. No clear hypothesis can be offered to suggest whether judged or behavior coordination might be the more effective measure.

Overview

Three studies were conducted using identical stimuli, similar procedures, and similar participants as judges. Results will be presented to test the following questions in order: (a) Are people reliable judges of coordination between partners? (b) Are judgments of "slices" of interaction predictive of behavioral coordination in whole interactions? (c) What behaviors predict judged coordination in the segments of interaction? (d) Does behavioral or judged coordination in segments predict interactants' attitudes from the whole conversation? The studies will be presented together rather than seriatim.

Method

Participants

In Study 1, participants were undergraduate students ($N = 51$) in a course on nonverbal communication who volunteered for a study on perceptions of interaction. Participants ranged in age from 18 to 26 (with the average at 20 years). All had English as a first or primary language. Twenty-three percent were African American or Hispanic; 17% identified themselves as Asian American. Eleven were born in countries other than the United States.

In Studies 2 and 3, participants were also undergraduates ($N = 17$ and 22, respectively) but from psychology courses at a large university in the West. They were demographically similar to the first sample.

Materials

For each of the three studies, 16 segments of videotaped interaction were selected for use. Selection of segments for inclusion proceeded as follows. Forty-nine 30-min interactions from a previous study (Cappella & Palmer, 1990) had been analyzed and coded for other purposes. Specifically, each had been evaluated for the degree of behavioral coordination over the entire 30-min conversation. Coordination was defined dynamically as the degree to which Person A's behavior was responsive to that of Person B's over and above A's prior behavioral response (Cappella, 1987). The operational procedures used to produce a measure of observed coordination for the interaction as a whole were based on time series analysis.[1] The coefficients from the time series analysis indicate the kind of coordination present: positive for reciprocal coordination of activity between partners; zero for the absence of coordination; negative for complementary coordination (high periods with low and low with high).

Coordination can occur on specific behaviors such as rate of speech (Webb, 1972; Street & Cappella, 1989) or on more global measures such as behavioral activity (Chapple, 1971). For the whole interactions, a global index of behavioral activity was used. The index was a z score sum of several behaviors during each 3-s interval of interaction: face-directed gaze, vocalization, illustrator gestures, adaptor gestures, and smiles and laughter.[2]

From the set of 49 interactions, 4 high in coordination and 4 low in coordination were selected for use in this study. Coordination scores for the interactions selected ranged from 0 to .66. The four low coordination ones had scores of 0, .19, 0, and .16, and the highs had scores of .66, .57, .54, .58 (see Table 1). The interactions selected for use here were either high or low in coordination across the entire interaction (30 min).

The four high- and four low-coordination interactions were balanced on sex composition with two male–male (MM) and 2 female–female (FF) conversations included. The interactants were all Caucasian and similar in age to the judges. Segments to be judged were chosen from these eight interactions by randomly selecting a 1-min slice from the first half and a 1-min slice from the second half. The procedure produced 16 1-min segments. The selection of segments was subject to the constraint that each interactant show some vocal participation. If one or the other of the partners was silent during the 1-min segment, the segment was rejected and another random starting point was selected.

In sum, a total of sixteen 1-min segments were chosen from eight different 30-min interactions, two per interaction from the first and second halves. Of the eight interactions used, four were high and four low in behavioral coordination. It is important to keep clear that when a randomly selected segment comes from a high (or low) coordination interaction, the segment itself may or may not be well coordinated behaviorally. The segments are not chosen to be high or low in coordination, the interactions from which the segments come are.

In Studies 1 and 2 the voices of the participants were audible. The content of what they were saying was not discernible because the voices were filtered through throat microphones that had the effect of cutting off high- and low-pitch sounds, effectively garbling the content. In Study 3 all vocal cues were eliminated by turning off the volume.

In Study 3, not only were vocal cues eliminated, but facial cues were reduced. The video channel was altered using the quantized mosaic technique pioneered by Berry (Berry, Kean, Misovich, & Baron, 1991; Berry, Misovich, Kean, & Baron, 1991) and used in coordination studies by Bernieri, Davis, Rosenthal, and Knee (1994). This procedure hides many of the details of facial action but does allow the audience to see movement in the mouth, eye, and brow regions of the face. Cues to specific facial emotions are removed, but facial animation resulting from vocalization and facial activity is visible.

The purpose of reducing vocal and facial cues in Study 3 was to evaluate whether a general form of movement coordination, independent of voice and facial emotion, was sufficient to activate judgments of coordination (Bernieri & Rosenthal, 1991). Because the voice and the face are the two most likely sources of information about emotion, removing them allows a test of just how much judged coordination in the presence of sources of emotional coordination (Studies 1 and 2) is due just to movement coordination (Study 3).

Design

The sixteen 1-min segments were shown to all participants. The segments represented a 2 × 2 × 2 × 2 within-subjects design, with the factors being coordination (high–low); sex composition of the dyad (MM–FF); position within the interaction (first half–second half); and replicate (Example 1–Example 2 of case). The order of presentation is presented in Table 1. In all three studies, all judges rated all 16 segments, so that coordination level, sex composition, position, and judge were fully crossed, whereas replicate was nested within factors but crossed with judges. ***Replicate*** here means that each of the eight conditions had two examples (e.g., Interactions 9 and 11 provided seg-

---

[1] Time series analysis assesses the degree of relationship between two series of values, in this case behavior scores. The relationship is expressed through a "transfer function model," which is very much like a regression of one series on the other series. The sign and magnitude of the regression-like coefficients indicate the type of coordination (positive for reciprocal and negative for complementary) and its degree (Cappella, 1996).

[2] Speaker–hearer role was covaried out of the index of coordination because speakers are much more active than listeners, and we wanted to know how much coordination in activity there is over and above that due to the complementarity of speaker and listener (see Cappella, 1996, for details).

Table 1
Means and Standard Deviations for Judged Coordination for 16 Segments: Position, Sex Composition, Coordination Condition, and Coordination Score for Three Studies

						Judged coordination					
						Study 1 (N = 51)		Study 2 (N = 17)		Study 3 (N = 22)	
Stim #/ Tape #	Replicate	Segment half	Sex comp	Coord magnitude	Coord condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
1/09	1	1st	FF	.00	Low	5.58	1.33	4.82	1.91	3.95	1.49
2/01	1	1st	MM	.19	Low	3.81	1.56	3.76	1.56	4.18	1.74
3/18	1	1st	FF	.66	High	6.47	1.19	6.00	.87	5.32	1.49
4/06	1	1st	MM	.57	High	6.68	.97	6.76	1.92	6.45	1.06
5/11	2	1st	FF	00	Low	7.01	1.28	6.71	1.83	6.36	2.38
6/05	2	1st	MM	.16	Low	5.22	1.41	4.82	2.10	5.18	1.68
7/02	2	1st	FF	.54	High	6.59	1.18	5.82	1.74	5.86	1.49
8/28	2	1st	MM	.59	High	5.77	1.66	5.17	2.43	5.82	1.87
9/11	1	2nd	FF	00	Low	6.61	1.17	6.53	2.51	5.27	1.35
10/05	1	2nd	MM	.16	Low	4.60	1.68	4.65	2.55	5.73	1.78
11/02	1	2nd	FF	.54	High	6.10	1.17	6.29	1.61	5.36	1.68
12/28	1	2nd	MM	.59	High	6.42	1.30	6.00	2.42	5.36	1.40
13/09	2	2nd	FF	00	Low	4.98	1.17	4.29	2.05	3.55	1.30
14/01	2	2nd	MM	.19	Low	5.56	1.57	5.24	2.39	5.64	1.92
15/18	2	2nd	FF	.66	High	4.95	1.52	3.18	1.55	4.04	1.25
16/06	2	2nd	MM	.57	High	6.11	1.31	6.65	1.37	5.73	1.32

Note. Comp = composition; coord = coordination; FF = female–female; MM = male–male. Judged coordination is the mean of four questions about coordination in Study 1 and of three questions in Studies 2 and 3. Coord magnitude is the degree of coordination over the entire interaction from time series coordination coefficient. Stim # is the order of appearance; tape # indicates the interaction in which the segment is found.

ments in the FF, low-coordination, first-half conditions—as well as the second-half condition).

There was only one order of presentation, with the segments rotating through the various conditions systematically. The first eight segments were from eight different interactions (first halves), and the second eight were from the second half of the same interactions (see Table 1).

Procedure

All the participants watched the edited video of segments in a large group. They were told that they would be watching and listening to segments from interactions between students like themselves and that they would "be asked to give their reactions to each segment immediately after it occurs." Each received a questionnaire requesting some demographic information, with each subsequent page posing questions about judged coordination for the video segment. Before seeing the segments and responding, they read definitions of criteria for judging the segments.

The video was shown on a large screen overhead projector easily visible to all. Each 1-min segment was followed by a screen with the words "Please Wait" and a 60-s pause while participants rated the segment on four (or three in Studies 2 and 3) questions concerning the partners' coordination.

Measures

Judged coordination. In Study 1, four measures of judged coordination were used. These were taken from Bernieri et al. (1988) and were posed in the form of 9-point scales with anchors of *very strongly agree* to *very strongly disagree.* The statements were as follows: "The partners engaged in simultaneous movement." "The partners had similar tempos of activity." "The partners' interaction was coordinated and smooth." "The partners matched one another's behaviors." These four measures were grouped into a single coordination measure based on the correlations among questions. The internal reliability of the four items was estimated for each of the 16 segments separately. The standardized alphas ranged from .63 to .80, with a mean of .72 and a standard deviation of 0.05. The four items were averaged to create an index labeled *judged coordination.*

In Studies 2 and 3, observers rated each of the segments on three aspects of coordination on a 9-point scale of *very strongly agree* to *very strongly disagree.* The rating scales were basically the same as the first three items used in Study 1, with the fourth item—a more static measure—dropped. The three items were averaged to create a judged coordination scale comparable to that of Study 1.

Behavioral coordination. Behavioral coordination was measured for the interaction as a whole and for each of the 1-min segments observed by the judges. The interactions from which segments had been selected had been previously coded for various vocal and kinesic behaviors. Included in this study are smiles, vocalization, face-directed gaze, adaptor gestures, and illustrative gestures. Coding rules and reliabilities are reported elsewhere (Cappella & Palmer, 1990). For the interaction as a whole, time series measures of coordination were used. (See Note 2 for a description of the measure.)

For the 1-min segments, two indexes were created for each behavior: an average amount of behavior for the pair and a (contemporaneous) correlation between the partners' behaviors defined over the time duration of the segment (twenty 3-s blocks). Ten indexes result for each segment: five mean behaviors and five correlations. The correlations are the measure of behavioral coordination for each segment.

Analysis

Analysis of results proceeded in three stages. First, hypotheses were tested on coordination within the interaction as a whole to see whether judges' ratings of coordination from brief segments predicted differences in behavioral coordination across the entire interaction. Next, specific behavioral patterns in each segment were related to judged coordination of the segments. Finally, the relative strength of behavioral and judged coordination was compared with the partners' attraction to one another and satisfaction with the conversation.

Results

Judged and Behavioral Coordination in the Entire Interaction

To test the hypothesis that judged coordination of the slices of interaction is related to the behavioral coordination of the interaction as a whole, two analyses were undertaken. First, judges' reliability with one another was assessed. A standardized Cronbach alpha among judges was .97, .85, and .86 across the three studies, respectively. The judges were sufficiently consistent with one another to warrant treating their mean scores as representative of the group. It should also be noted that the judges for Study 3 were no less reliable than those in Study 2, despite the lower amount of information available in face and voice.

These means across judges are presented in Table 1 along with other information about the study's design. The means for judged coordination of the 16 segments correlate strongly across the three studies. Study 1 and Study 2 correlate at .88; Studies 2 and 3, at .75; Studies 1 and 3, at .64. All are significant at $p < .007$, at least, and do not differ by chi-square test for difference among correlations, $\chi^2(2, N = 16) = 1.79$ (Hunter, Schmidt, & Jackson, 1982). These strong positive correlations suggest that despite differences in sample size, measures of judged coordination, and even differences in visual and auditory information, means across judges are robust indicators of coordination.

Studies 2 and 3 differ only in the presence of cues to vocal and facial emotion. They share 53.3% of their variance (adjusted R^2). Even Studies 3 and 1 share 37.3% of their variance in judged coordination, differing in both questions used to judge coordination and facial and vocal cues to emotion. The amount of shared variance between Study 3 and the others suggests that a substantial portion of judged coordination is due to behavioral cues other than facial and vocal emotion.

A within-subjects analysis of variance (ANOVA) was run on the judged coordination scores. Each participant judged 16 segments that varied by position in the interaction (first half–second half), sex composition (MM or FF), coordination level (high or low), and example. The within-subjects results from multivariate analysis of variance (MANOVA) are presented in Table 2. Two conclusions can be drawn from Tables 1 and 2. First, there is a significant main effect in all three studies such that the interactions low in behavioral coordination are low in judged coordination: $M(\text{low}) = 5.42$ and $M(\text{high}) = 6.14$ in Study 1; $M(\text{low}) = 5.10$ and $M(\text{high}) = 5.73$ in Study 2; and $M(\text{low}) = 4.98$ and $M(\text{high}) = 5.49$ in Study 3.

In Study 1, the main effect is quite strong, $F(1, 50) = 79.52$, $p < .001$. In Studies 2 and 3 it is less powerful but still significant. These main effects in each study are embedded in complex higher order interactions with sex composition of the dyad, segment of the interaction, and example. These interactions, some strong, indicate that judged coordination is sensitive to

Table 2
F and p Values for Judged Coordination: Main and Interaction Effects From Within-Subjects MANOVA for Three Studies

	Study 1		Study 2		Study 3	
Effect	$F(1, 50)$	p	$F(1, 16)$	p	$F(1, 21)$	p
Segment	10.01	.003	.53	.48	3.12	.09
Example	.02	.90	3.66	.07	.15	.70
Coord	79.52	.000	4.86	.04	6.22	.02
Sex comp	29.09	.000	.11	.74	6.05	.02
Seg × Ex	35.05	.000	6.73	.02	13.96	.001
Seg × Coord	10.84	.002	1.84	.19	4.95	.04
Seg × Sex	49.53	.000	8.10	.01	23.01	.000
Ex × Coord	44.79	.000	13.72	.002	6.49	.02
Ex × Sex	12.08	.001	6.86	.02	.60	.45
Coord × Sex	68.91	.000	12.71	.003	.80	.38
Seg × Ex × Coord	18.14	.000	5.94	.03	21.50	.000
Seg × Ex × Sex	22.98	.000	19.21	.000	16.01	.001
Seg × Coord × Sex	1.16	.287	.53	.48	10.35	.004
Ex × Coord × Sex	16.85	.000	.11	.74	.09	.77
Ex × Coord × Sex × Seg	2.50	.12	1.3	.27	.02	.89

Note. MANOVA = multivariate analysis of variance; coord = coordination; Seg = segment or piece of video from first or second half of interaction; Ex = example or replication of segment type in terms of position, coordination level, and sex of interactants; Coord = coordination of the behaviors, high or low; Sex comp or sex = sex composition of the interactants, male–male or female–female.

the particular slice being evaluated and not just to the coordination level of the interaction from which the slice comes.

However, a case can be made that the nature of the interactions is not as complex as it might first seem. Clark (1973) has argued that treating replicates as fixed effects rather than as random effects can lead to excessively liberal hypothesis tests, which in the present case could overstate not only substantive tests of hypotheses about coordination but also overstate the significance of the higher order interactions. Whether the replicates in this design should be treated as fixed or random is disputable (Cohen, 1976; Keppel, 1976; Wilke & Church, 1976).

The most balanced description of the results is represented in Figure 1, in which the interaction between sex composition and behavioral coordination is displayed for all three studies. The reasoning is as follows. Clark's approach to the analysis, which requires indeterminate quasi-F ratios and imprecise degrees of freedom, is more complex than necessary. It is also very conservative. Because persons are reliable judges of stimuli for each study (as we have already seen), the average score across judges is an acceptable measure. Averaging across persons yields a design in which examples are nested within levels of coordination by sex composition by segment. Examples are then treated as a random factor, with the other three treated as fixed factors.[3]

[3] The two ways of analyzing the data are liberal in the case of treating example as fixed and conservative when judge's scores are averaged

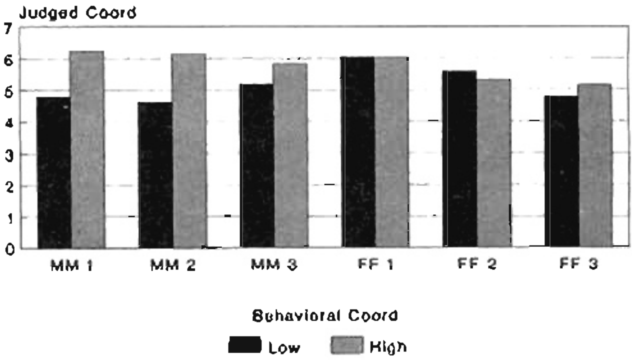

Figure 1. Mean judged coordination for sex composition of the dyad and behavioral coordination level: Studies 1, 2, and 3. Coord = coordination; MM = male–male; FF = female–female.

A multivariate test of judged coordination from Studies 1, 2, and 3 on coordination, sex composition, and segment, with replicate nested within levels, yields only one significant multivariate effect: sex composition by coordination level, Pillai's = .502, $F(3, 10) = 3.36$, $p = .06$. The univariate *F*s in each case are as follows: for Study 1, $F(1, 12) = 3.94$, $p = .07$; for Study 2, $F(1, 12) = 2.88$, $p = .11$; for Study 3, $F(1, 12) < 1$, *n.s.* The means are displayed in Figure 1 and show clearly that differences from high to low coordination are readily judged in male–male dyads and not well judged in female–female dyads.

Students interviewed informally after the experiment denied that they were able to make judgments of coordination with consistency, and they disavowed any confidence in their own ratings. Yet, the data show that untrained observers can take small slices of interaction and reach reliable conclusions about the extent of behavioral coordination at least for male dyads. Even when the behavioral coordination is assessed on the interaction as a whole, judges can sense lower and higher levels of coordination in samples from the whole, especially for male–male interactions.

Behavioral and Judged Coordination in the 1-Min Segments

In evaluating the relationship between judged and behavioral coordination, level of analysis shifts. Behavioral coordination in each segment is defined as the correlation between Partner A's behavior and Partner B's in the twenty 3-s units of each segment. For example, in Segment 1 the correlation between A's and B's smiles across the 20 time units is the measure of behavioral coordination on smiles for that segment. In the same segment, the average duration of smiling for the pair is also calculated. The shift in levels of analysis is necessary to properly evaluate the link between judged and behavioral coordination.

over judges and examples are treated as a random variable. The only robust finding is the interaction between sex composition and coordination level.

Correlates of judgments. Table 3 reports the correlations between behavioral indexes of coordination and judged coordination for all three studies. Ten behavioral measures are listed: static measures on five behaviors and dynamic measures on five behaviors. The static measures do not represent patterns of coordination and are presented here only for comparison and later control. Only a few correlations reach standard levels of statistical significance, mostly because of the low power from having only 16 segments. None of the correlations reach standard levels of significance in predicting the modified video segments used in Study 3.

Nevertheless, certain patterns deserve comment and further exploration. Coordination in partners' smiles is positively correlated with judged coordination in all three studies, with two of the three correlations significant. This pattern suggests that when partners' smiles are on or off in concert, observers judge the partners to be coordinated. No other behavior shows a similar effect.

The coordination between partners on illustrators and on gaze shows the opposite correlation to judged coordination. Consider gaze and gaze aversion. The negative correlation suggests that when gaze by Person A goes with gaze aversion by the partner, then judged coordination is enhanced. Similarly, when speech-related gestures covary temporally with the absence of such gestures by the partner, judged coordination is higher. A similar but nonsignificant trend is found with adaptor gestures.

Coordination on behaviors associated with speaker and listener role (gestures, gaze) may affect judged coordination differently from behaviors related to emotional state (e.g., smiles).

Table 3
Correlations, Mean Correlations, and Chi-Square Differences Between Judged Coordination and Behavioral Coordination in Three Studies

	Judged coordination				
Variable	Study 1	Study 2	Study 3	Mean r	$\chi^2(2, N = 16)$
Means					
Smiles	.42*	.38	.30	.37	0.15
Illus	−.10	.14	.30	.13	1.66
Vocal	−.07	.22	.35	.17	1.56
Gaze	−.10	.10	.04	.01	0.34
Adaptors	.42*	.20	.30	.31	0.49
Correlations					
Smiles	.64***	.59**	.32	.52	1.70
Illus	−.39	−.58**	−.26	−.41	1.18
Vocal	.08	−.15	.04	0	0.49
Gaze	−.54**	−.33	−.11	−.32	1.82
Adaptors	−.32	−.23	−.16	−.24	0.23
Coord condition[a]	.44*	.30	.31	.35	0.25
Coord score	.27	.15	.29	.24	0.19

Note. Correlation = correlation between partners' behaviors in the segment; Mean = average duration of partners' behaviors in the segment; Illus = illustrative gestures; Coord condition = 1,0 value for coordination of the interaction as a whole; Coord score = the coordination score for the whole interaction.
[a] High = 1, low = 0.
* $p < .10$. ** $p < .05$. *** $p < .01$. All tests are two-tailed.

When behaviors indicative of speaker and listener role occur in a complementary fashion, then the partners are avoiding speaker–hearer role conflict. For example, if both are gesturing at the same time, then both are trying to occupy the speaker role. If both are gazing at the partner at the same time, then both are exhibiting behaviors associated with the listener role at the same time. Observers may be implicitly sensitive to such politeness norms and may judge partners who violate them as less coordinated with one another.

Smiling is not associated with occupying the speaker or hearer role, and so it does not function in the same way. Smiling is more typically an indicator of positive affect (or at least an attempt to control another's hostile affect). When positive facial affect by one person is associated with positive facial affect by the other (and negative with negative), then the correlation appears to yield judgments of greater coordination between the partners.

In short, our first, tentative conclusions about behavioral and judged coordination are that (a) judged coordination results from positive behavioral coordination on indicators of affect and (b) judged coordination results from complementarity (negative contemporaneous correlation) on indicators of speaker–hearer role.

Confounding effects of mean behaviors. The first potential confound to consider is the mean behavioral level. It is possible that coordination in smiling will not predict judged coordination when mean levels of smiling are controlled. Judges may simply rate smilers higher in coordination regardless of the patterning of smiles during interaction. A similar confound could operate with behaviors associated with speaker–hearer role.

To test this possibility, regression analyses were run on judged coordination, with two predictor variables (e.g., smile coordination and mean smiles) for each of the five behaviors. Both mean and correlation were entered at the same time, but only one behavior was used as a predictor in this analysis. The results appear in Table 4. The magnitude, sign, and level of statistical significance for each of the behavioral coordination coefficients are very similar to the zero-order correlations. This means that behavioral coordination still predicts judged coordination even after the mean amount of behavior is controlled. Judges are not simply imputing synchrony to partners who talk more, smile more, or gesture more. They seem to actually use the timing of the behavioral actions in their judgments. As before, no indicator of behavioral coordination reaches standard levels of significance in predicting the no-voice, mosaic video segments.

Table 4
Betas From Regressions of Individual Behaviors (Partner Correlations and Mean) on Judged Coordination in Three Studies

	Judged coordination					
	Study 1		Study 2		Study 3	
Behavior	β	M	β	M	β	M
Smiles	.57***	.26	.53***	.25	.26	.24
Illustrate	−.41	−.15	−.57**	.07	−.23	.28
Vocals	.06	−.05	−.10	.20	.16	.40
Gaze	−.53**	−.03	−.35	.15	−.12	.06
Adaptors	−.30	.41	−.22	.19	−.15	.27

Note. β = correlation between partners' behaviors in the segment; M = mean duration of partners' behaviors in the judged segment; Illustrate = illustrative gestures.
** $p < .05$. *** $p < .01$.

Table 5
Betas, p Values, and Adjusted R^2 From Stepwise Regressions of Behavioral Predictors of Judged Coordination in Three Studies

	Judged coordination					
	Study 1		Study 2		Study 3	
Variables	β	p	β	p	β	p
Means						
Smile					.35	.20
Illus					.42	.12
Gaze						
Adaptor						
Correlations						
Smile	.53	.02	.43	.06	.27	.29
Illus			−.42	.07		
Gaze	−.39	.07				
Adaptor						
Adj R^2	.47	.01	.42	.01	.15	.20

Note. Step regressions for Studies 1 and 2 used a p-in criterion of .15 and a p-out criterion of .20; Study 3 used .25 and .30, respectively. Illus = illustrative gestures; Adaptor = adaptor gestures; Correlation = correlation between partners' behaviors in judged sample; Mean = average duration of partners' behaviors in the judged sample.

The behaviors that account for judged coordination in Studies 1 and 2 are coordination on smiles and coordination on either gaze (Study 1) or illustrative gestures (Study 2). Judges seem to treat partners who smile in concert and who gesture or gaze in complementary fashion as more coordinated. The same sign patterns exist in Study 3 but not at normal levels of statistical significance. Controlling for mean levels of behavior does not alter these effects.

Multiple indicators. Individual correlations between behavioral measures and judged coordination do not tell us how much overlap there is between the two sets. To maximize the variance in judged coordination accounted for by behavioral indicators, a stepwise regression procedure was used.[4] Eight behavioral indicators were included as potential predictors of judged coordination: the mean and correlational indexes for smiles, illustrators, gaze, and adaptors. The results are presented in Table 5. The only values listed in the table are those that reach preassigned levels of inclusion and exclusion. The implications of the results from the stepwise regression are similar to the findings from taking one behavior at a time.

For Study 1, partners' coordination on smiles predicts judged

[4] The vocalization behavior had to be dropped from the stepwise analyses because of excessive collinearity with the other predictors. Because it shows the lowest levels of correlation with judged coordination anyway, this was not a serious loss.

coordination positively, whereas their correlation on gaze predicts coordination negatively. Approximately 47% of the variance in judged coordination is explained by these two behavioral indicators of coordination. For Study 2, approximately 42% of the variance in judged coordination is explained by partners' behavioral coordination on smiles and on illustrative gestures. As before, Study 3 produces no statistically significant predictors of judged coordination, explaining a nonsignificant 14% of the variance.

Accounting for Study 3 judgments. The behavioral basis of judged coordination in Studies 1 and 2 shows a reasonable amount of consistency. Despite the high correlation in judged coordination between Study 3 and its predecessors, the five behaviors studied do not account for appreciable variance in judgments. This is something of a paradox. If there were low correlations between Study 3 and the other two studies, then one might argue that the absence of vocal or clear facial cues led to incoherent judgments; however, the judgments share considerable variance across studies.

One argument is that the judges are using the same behaviors in the same way in making their ratings in Study 3, but the influence of behaviors is reduced because of the procedures of Study 3. The mosaic technique, after all, does not remove facial animation as a cue, it simply masks it. In support of this contention, the pattern of signs on indexes of behavioral coordination from Study 1 to Study 3 and from Study 2 to Study 3 (see Table 4) is too consistent to be a chance occurrence. Nine of the 10 correlations have the same sign, which is highly significant by a test for proportions ($z = 26, p < .001$). In addition, the correlations across the three studies do not differ significantly by a chi-square test (Hunter et al., 1982). Table 3 presents the average correlation as well as the chi-square measures of variation in the correlations across the three studies. None of the differences approach statistical significance.

The foregoing results support a conclusion that judgments in all three studies are being made in basically the same way. Observers are looking for positive covariation in the partners' facial activity and for negative covariation in the partners' floor-holding behaviors as signs of coordinated activity. The pattern of signs and the absence of significant differences from Study 1 through Study 3 in correlations between behavioral and judged coordination suggest that judgments are similar across studies and simply noisier in Study 3. One possibility of course is that the absence of affective information from face or voice in Study 3 is a significant loss in judgments of coordination.

Predicting Participants' Ratings

Shared variance is only one criterion of success in linking judged and behavioral coordination. Predicting interpersonal and conversational outcomes may be more important. In the study from which stimuli were selected, partners rated their satisfaction with the conversation, and their attitudes toward the partner.[5]

Table 6 presents the correlations between partners' attitudes and both judged and behavioral coordination. The table includes a judged coordination index based on all three studies[6] and a coordination score that is a measure of behavioral coordination for the entire 30 min interaction, not just the 1-min segments.

Table 6
Correlations Between Objective Coordination, Judged Coordination, and Participants' Reported Dissatisfaction With the Conversation and Attraction to the Partner

Coordination measure	Conversational dissatisfaction	Attitude toward partner
Judged coord (Mean, 3 studies)	.47**	.41
Coordination score[a]	−.47**	.26
Correlations		
Smile	−.08	.59*
Illus	.08	−.06
Vocal	−.15	.18
Gaze	.21	.07
Adaptor	−.14	−.33

Note. Coord = coordination; Illus = illustrative gestures; Adaptor = adaptor gestures.
[a] Coordination score is a measure of behavioral coordination obtained across the entire interaction.
* $p \leq .05$. ** $p \leq .10$. All tests are two-tailed.

The self-reports of conversational dissatisfaction and attitude toward the partner are based, of course, on the entire conversation, whereas the other measures of judged coordination and the individual behavioral correlations are based on 1-min slices.

Attitude toward the partner is positively and significantly related to coordinated smiling in the 1-min segments. Judged coordination has a moderate but nonsignificant correlation to attitude toward the partner. In a stepwise regression using all the descriptors of the 1-min segments, only coordinated smiling accounts for variance in attitude toward the partner, $F(1, 14) = 7.40, p = .017$, adjusted $R^2 = .30$. No other predictor is close to significance. Controlling for mean level of smiling in the segments does not alter the results at all. Of interest is the fact that coordinated smiling in the 1-min segments predicts attitude toward the partner better than do coordination on a variety of behaviors across 30 min or judged coordination for the 1-min segments.

Conversation dissatisfaction shows no statistically significant relationships with any predictors, although judged coordination and the overall coordination are close. What is odd is the direction of the relationship—the more judged coordination present, the more dissatisfied partners are with the conversation. In a stepwise regression of all descriptors of the 1-min segments, two variables account for 30% of the variance, $F(2, 13) = 4.20, p = .039$, adjusted $R^2 = .30$, in dissatisfaction with the conversation: judged coordination, $\beta = .753, t(13) = 2.87, p = .013$, and smile coordination, $\beta = -.508, t(13) = 1.94, p$

[5] Conversational dissatisfaction was measured by questions about not feeling involved, insufficient opportunity to participate, and the feeling that the "back and forth flow" was inadequate. Attitude toward the partner included liking, the desire to work with the partner again, and perceiving the partner to be well adjusted (Byrne, 1971).

[6] The correlations between judged coordination and attitude toward the partner are .47 (Study 1), .39 (Study 2), and .26 (Study 3). For conversational satisfaction, the individual correlations are .29, .49, and .49, respectively. Because the signs and magnitudes are comparable, an index averaging the three was created.

= .075. These effects are unchanged when mean level of smiling is controlled. This is a classic example of a suppressor effect (Cohen & Cohen, 1983, pp. 95–96) because judged coordination and coordination in smiles correlate at r = .565, p = .023. Once smile coordination enters as a predictor, the direct effect for both smile coordination and judged coordination become stronger. Coordinated smiling is associated with greater satisfaction with the quality and character of the conversation, whereas judged coordination is associated with more dissatisfaction with the conversation.

Coordinated smiling in the segments accounted for significant variance in overall attraction between partners and in satisfaction with the conversation. Judged coordination in the segments did not add to variance explained in partners' attraction over that due to coordinated smiling. Although judged coordination for segments accounted for significant variance in partners' conversational satisfaction, the relationship was surprisingly a negative one. Coordinated smiling was positively related to partners' conversational satisfaction, and it added approximately half of the 30% variance explained by the two factors.

Conclusions and Implications

These studies were undertaken to understand how to best measure coordination in human interaction—through judgments of naive raters or through behavioral coding. The findings can be summarized by focusing first on reliability and consistency, next by judging whole interactions from "slices," and finally by considering the specific relationship between coded and judged segments.

Can Untrained People Make Judgments of Coordination?

In all three studies, judges were consistent with one another. The internal reliability of the larger and smaller groups, those with full information and partial information, was quite good. This finding provides additional support for Bernieri's previous work (Bernieri, 1988; Bernieri, Gillis, Davis, & Grahe, 1996; Bernieri et al., 1988, 1994) and suggests that untrained judges can generate sound measures of the synchrony between partners in brief segments of conversation.

Not only are people able to make these judgments consistently, but the rank ordering of segments on mean judged coordination is quite consistent from one group to the next. The three studies conducted here differed in the kinds of questions used to rate coordination and in the amount of information available to the judges. No voice or cues to facial emotion were available in the third study. Despite these differences correlations among mean ratings for segments were quite strong. The reliability across judges within study and the consistency across studies in judged coordination suggests that the judgment method developed by Bernieri is a robust one. But what does it measure?

Wholes and Slices

The eight interactions chosen for study were originally 30 min long. They were evaluated as high or low in coordination on a time series index of behavioral activity (including gaze, gesture [2 types], smiles, and vocalization). The index used all 30 min of interaction. One very stringent test of the utility of the judged coordination measure is whether judged coordination of segments of an interaction can replace measures of behavioral coordination for the entire interaction. If judged coordination of an interactional segment is a good predictor of behavioral coordination in the entire interaction, then considerable efficiencies arise.

Ambady and Rosenthal (1992) concluded that judgments of small slices of interaction were often good indicators of wholes, but the study reported here moves Ambady and Rosenthal's question from the level of the individual to the level of the dyad. Our data show that judged coordination of segments of interaction from male–male dyads is associated with the overall behavioral coordination of those dyads. But the same claim cannot be made of female–female dyads. One explanation for the difference between the two is found in the fact that the female segments differed from the male interactions in the level of coordinated smiling. That is, the segments of male–male interactions that were chosen from the high-coordination conversations also had elevated levels of coordinated smiling. The correlation between overall coordination score and the level of coordinated smiling of the segment was .61 for the male dyads. For the female dyads, the correlation was actually negative, −.26. Some of the high-coordination female–female segments actually had low levels of coordinated smiling and vice versa.

In addition, judged coordination averaged across the three studies correlated .57 with smile coordination. For female–female dyads, judged coordination relied on smile coordination very strongly, r = .85; for male–male dyads, r = .36.

So there are two reasons for the absence of relationship between judged coordination of segments and overall behavioral coordination for the female–female examples. First, judges were evaluating the female partners' coordination strongly in terms of whether they smiled in synchrony or not. Second, through bad luck, the segments of female interaction that were behaviorally coordinated overall were not coordinated in partners' smiling. This explanation is not just a methodological problem but is a substantive conclusion. It shows that judged coordination is dependent on cues associated with mutual smiling. People appear to be sensitive to this form of behavioral contagion during interaction (Hatfield, Cacioppo, & Rapson, 1994) and are capable of recognizing it, treating it and its absence—even from brief slices of interaction—as evidence of overall coordination.

Which Behaviors Yield Judged Coordination?

To determine just what the bases of judged coordination were in behavioral terms, our analysis moved to the level of the segments. Behavioral coordination, indexed by correlations between the partner's individual behaviors (gaze, illustrator gestures, adaptors, vocalization, and smiles), was measured for each 1-min segment. Behavioral and judged coordination shared from 40% to 50% of their variance when only three interactional variables were considered: smiles, gaze, and illustrative gestures. This substantial amount of shared variance suggests that people are sensitive to behavioral coordination when they view and hear interactions, their protests to the contrary. It also suggests that judgments of coordination are not an artifact of spurious third variables (such as the mean levels of gesture, gaze, etc.) but in fact directly tap into behavioral patterns of synchrony.

The data also yield some clues about what it is that judges are focusing on when they rate coordination. When rating interactions for coordination, judges are probably looking for signs of simultaneity of affect (e.g., as manifested by simultaneous facial emotion and perhaps more generally by simultaneous facial animation). To verify these inferences, we would need to show that affect in other outlets (e.g., in the voice or in body postures) also produces judgments of coordination or that nonsimultaneous facial affect (e.g., smiles that are complementary in timing) yields judgments of noncoordination. Given the importance of affect in the human face in the development of attachment to the primary caregiver (Malatesta, Culver, Tesman, & Shephard, 1989) and the importance of the timing of facial responsiveness between infant and caregiver (Schaffer, Collis, & Parsons, 1977), it is not surprising that synchrony in facial affect would be treated as if it were an important cue in judgments of coordinated interactions.

Judges also seem to use behavioral signs of speaker and hearer role in their ratings of coordination. Specifically, when face-directed gaze or illustrative gestures are complementary, judged coordination tends to be higher. If these behaviors are treated as signs of speaker–hearer role, then our evaluators may be judging stereotypically polite interactions as more coordinated. When speaker and hearer roles do not overlap—the stereotype, although not necessarily the reality, of polite interaction—gestures do not overlap either and gaze by one is accompanied by gaze aversion by the partner. In this scenario, partners' behaviors are judged to be meshed in the sense that neither is trespassing on the conversational space of the partner.

An interesting test of the limits of this explanation would use an interaction in which both partners are highly involved so that their turns are consistently overlapping. Although such an interaction is stereotypically impolite, it may still be judged as coordinated. The point is that even though norms of politeness exist in a variety of cultures (Brown & Levinson, 1987), their form may vary considerably from culture to culture. Turn-taking norms are one such example, varying by ethnicity and race. We do not know whether judges from different ethnic groups would use behavioral cues of speaker–hearer role differentiation differently.

Our data do not simply confirm again Bernieri's (1988; Bernieri et al., 1988) findings that judgments of coordination are reliable but extend their validity. Accepted measures of behavioral coordination covary with judged coordination and, moreover, do so in a way that the behavioral basis of judged coordination is now clearer. The studies presented here imply that judged coordination is based on synchronous affect and complementary signs of speaker–hearer role.

Some have argued that partners coordinate their activities in social interaction in ways that can only be perceived as gestalt units. The claim, in effect, is that coordination does not reside in any particular behavior but in some social affordance captured by the behaviors operating together (Baron & Boudreau, 1987; Bernieri & Rosenthal, 1991; Newtson, 1994). The data presented here indicate that the correlations between specific behaviors are associated with perceptions of partner coordination. However, I do not think that the findings here are necessarily inconsistent with gestalt theories of the perception of coordination. After all, over half of the variance in judged coordination remains to be explained. Gestalt measures of behavioral coordination may be found to be more successful than particular measures, and it would be odd indeed if micromeasures of behavioral coordination were found to be unrelated to gestalt measures.

Interactional Outcomes

Both judged and behavioral coordination successfully predicted outcome ratings by interactional partners. The fact that the two measures of coordination showed any relationship to participants' evaluations of one another or their evaluations of the interaction is rather remarkable. After all, the partners' self-reports were based on full information (including verbal discourse) and on 30 min of conversation. The judges' ratings and the behavioral coordination scores were based on single minutes of interaction from the whole. The presence of any correlation between interaction pattern (judged or coded) and outcome must be considered significant testimony to the importance of coordination to attraction and conversational satisfaction.

Synchronous smiling was again implicated as a key predictor. The stronger the synchronous smiling, the more positive the partners' attitudes toward one another, and the result was similar for satisfaction with the quality of the conversation. Not only do judges of interactions rely on mutual smiling but participants in the interactions themselves may as well. Of course, the direction of causality is not clear. Positive attitudes about the partner and the conversation may produce more coordinated smiling.

The truly perplexing result is that judged coordination was related to dissatisfaction with the conversation.[7] To my knowledge, no other research using the judged coordination measure has tested its relationship to partners' conversational satisfaction. So this finding does not contradict any previous work. One possibility is that judged coordination taps into a kind of stereotype of polite interaction, perhaps even "hyperpoliteness." Partners who rigidly adhere to rules of turn-taking, not overlapping one another's conversational space, carefully remaining in the listener role until the speaker gives up the floor, may be judged to be coordinated conversationally. This characterization is consistent with the findings presented here on judged coordination and complementarity in gaze and gesture. But such hyperpoliteness may not be experienced by interactants as involving or satisfying in terms of conversational back and forth. Observers may judge rigid turn-taking as coordinated, but participants may experience it as less than satisfying.

Testing this explanation requires having interactional segments in which partners overlap one another's turns regularly and others where they do not. If the interactions can be selected from populations where conversational norms differ appreciably, so that conversational intrusion is the norm in one group

[7] Partialling out various factors such as sex composition, other indicators of coordination in the segment, mean levels of behavior, and total coordination score not only does not depress the effect but actually enhances it somewhat.

but not the other, then a clear test would be possible. Judged coordination should be low in interactions where overlap occurred, but the experience of overlapped and nonoverlapped turns would differ by group norm. Such a test remains for future research.

Which Measure?

The data from the studies reported in this article hold out real hope that coordination in social interaction can be studied using judgment methods and slices of interaction rather than behavioral coding of lengthy interactions. Before interaction researchers allow granting agencies to use the line item veto for their coding budgets, several lines of further study should be explored.

Over half of the variance between coding and judgment measures is still not explained. Our analyses used only a few behaviors, none of them verbal ones, and they ignored potentially important aspects of bodily movement and vocal affect. Too, variation among judges has not been studied. When judges are instruments, then variation among instruments must be a part of the process of measurement evaluation. Important variation across populations and settings must also be completed before similarity between judgment and coding procedures can be pronounced successful. For example, should judges and interactants match demographically? Will similar patterns of behavioral coordination in infants, toddlers, and children be associated with judgments of coordination? Despite these remaining questions, researchers can now seriously entertain the possibility that judged coordination is a surrogate for behavioral coordination.

References

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274.

Baron, R. M., Amazeen, P. G., & Beek, P. J. (1994). Local and global dynamics of social relations. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in psychology* (pp. 111–138). San Diego, CA: Academic Press.

Baron, R. M., & Boudreau, L. A. (1987). An ecological perspective on integrating personality and social psychology. *Journal of Personality and Social Psychology, 53,* 1222–1228.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher–student interactions. *Journal of Nonverbal Behavior, 12,* 120–138.

Bernieri, F. J., Davis, J. M., Rosenthal, R., & Knee, C. R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of sound and facial affect. *Personality and Social Psychology Bulletin, 20,* 303–311.

Bernieri, F. J., Gillis, J. S., Davis, J. M., & Grahe, J. E. (1996). Dyad rapport and the accuracy of its judgment across situations: A lens model analysis. *Journal of Personality and Social Psychology, 71,* 110–129.

Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother–infant interactions. *Journal of Personality and Social Psychology, 54,* 243–253.

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 401–432). Cambridge, England: Cambridge University Press.

Berry, D. S., Kean, K. J., Misovich, S. J., & Baron, R. M. (1991). Quantized displays of human movement: A methodological alternative to the point-light display. *Journal of Nonverbal Behavior, 15,* 81–97.

Berry, D. S., Misovich, S. J., Kean, K. J., & Baron, R. M. (1991). Effects of disruption of structure and motion on perceptions of social causality. *Personality and Social Psychology Bulletin, 18,* 237–244.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language use.* Cambridge, England: Cambridge University Press.

Byrne, D. (1971). *The attraction paradigm.* New York: Academic Press.

Cappella, J. N. (1987). Interpersonal communication: Fundamental questions and issues. In C. R. Berger & S. Chaffee (Eds.), *The handbook of communication science* (pp. 184–238). Beverly Hills, CA: Sage.

Cappella, J. N. (1994). The management of conversational interaction in adults and infants. In M. L. Knapp & G. R. Miller (Eds.), *Handbook of interpersonal communication* (2nd ed.; pp. 380–419). Newbury Park, CA: Sage.

Cappella, J. N. (1996). Dynamic coordination of vocal and kinesic behavior in dyadic interaction: Methods, problems, and interpersonal outcomes. In J. Watt & C. A. Van Lear (Eds.), *Dynamic patterns in communication processes* (pp. 353–386). Thousand Oaks, CA: Sage.

Cappella, J. N., & Flagg, M. E. (1992, July). *Interactional adaptation, expressiveness and attraction: Kinesic and vocal responsiveness patterns in initial liking.* Paper presented at the International Society for the Study of Social and Personal Relationships, Orono, ME.

Cappella, J. N., & Palmer, M. T. (1990). Attitude similarity, relational history, and attraction: The mediating effects of kinesic and vocal behaviors. *Communication Monographs, 57,* 161–183.

Cappella, J. N., Palmer, M. T., & Donzella, B. (1991, May). *Individual consistency in temporal adaptations in nonverbal behavior in dyadic conversations: High and low expressive groups.* Paper presented at the International Communication Association Conference, Chicago.

Cappella, J. N., & Street, R. L. (1985). A functional approach to the structure of communicative behavior. In R. L. Street & J. N. Cappella (Eds.), *Sequence and pattern in communicative behavior* (pp. 1–29). London: Edward Arnold.

Chapple, E. D. (1940). Measuring human relations: An introduction to the study of interaction of individuals. *Genetic Psychology Monographs, 22,* 3–147.

Chapple, E. D. (1971). Toward a mathematical model of interaction: Some preliminary considerations. In P. Key (Ed.), *Explorations in mathematical anthropology* (pp. 141–178). Cambridge, MA: MIT Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Cohen, J. (1976). Random means random. *Journal of Verbal Learning and Verbal Behavior, 15,* 261–262.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Duncan, S., & Fiske, D. W. (1977). *Face-to-face interaction.* Hillsdale, NJ: Erlbaum.

Gottman, J. M. (1979). *Marital interaction.* New York: Academic Press.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion.* New York: Cambridge University Press.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Isabella, R. A., & Belsky, J. (1991). Interactional synchrony and the origins of infant–mother attachment: A replication study. *Child Development, 62,* 373–384.

Isabella, R. A., Belsky, J., & van Eye, A. (1989). Origins of mother–infant attachment: An examination of interactional synchrony during the infant's first year. *Developmental Psychology, 25,* 12–21.

Keppel, G. (1976). Words as random variables. *Journal of Verbal Learning and Verbal Behavior, 15,* 263–265.

Malatesta, C. Z., Culver, C., Tesman, J. R., & Shephard, B. (1989). The development of emotion expression during the first two years of life. *Monographs for the Society for Research in Child Development, 54* (1&2, Serial No. 219).

Margolin, G., & Wampold, B. E. (1981). Sequential analysis of conflict and accord in distressed and nondistressed marital partners. *Journal of Consulting and Clinical Psychology, 49,* 554–567.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago: University of Chicago Press.

Newtson, D. (1994). The perception of the coupling of behavior waves. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in psychology* (pp. 139–167). San Diego, CA: Academic Press.

Palmer, M. T. (1989). Controlling conversations: Turns, topics, and interpersonal control. *Communication Monographs, 56,* 1–18.

Patterson, M. L. (1983). *Nonverbal behavior: A functional perspective.* New York: Springer-Verlag.

Pike, G. R., & Sillars, A. L. (1985). Reciprocity and marital communication. *Journal of Personal and Social Relationships, 2,* 303–324.

Rosenthal, R. (1987). *Judgment studies: Design, analysis and meta-analysis.* Cambridge, England: Cambridge University Press.

Schaffer, H. R., Collis, G. M., & Parsons, G. (1977). Verbal interchange and visual regard in verbal and preverbal children. In H. R. Schaffer (Ed.), *Studies in mother–infant interaction* (pp. 291–324). New York: Academic Press.

Street, R. L., & Cappella, J. N. (1989). Social and linguistic factors influencing adaptation in children's speech. *Journal of Psycholinguistic Research, 18,* 497–519.

Tickle-Degnen, L., & Rosenthal, R. (1987). Group rapport and nonverbal behavior. *Review of Personality and Social Psychology, 9,* 113–136.

van Lear, C. A. (1991). Testing a cyclical model of communicative openness in relationship development: Two longitudinal studies. *Communication Monographs, 58,* 337–361.

Webb, J. T. (1972). Interview synchrony: An investigation of two speech rate measures. In A. W. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 115–133). New York: Pergamon.

Wilke, E. L., & Church, J. D. (1976). Comments on Clark's "The language-as-fixed-effect fallacy." *Journal of Verbal Learning and Verbal Behavior, 15,* 249–255.

Received January 25, 1995
Revision Received June 20, 1996
Accepted August 15, 1996 ■